From space to place: place-based explorations of text

Purves, Ross S ; Derungs, Curdin

Abstract: New data sources, for example in the form of geotagged image libraries and digitised archives of historical text documents, provide us with new opportunities for exploring how place is described. Using a framework derived from work in human geography and information science, we illustrate how there is more to place than names and coordinates. Through a set of case studies we explore different aspects of the seemingly trivial query 'mountains in the Alps' addressing a range of issues including ambiguity, the use of vernacular names, ways in which concepts such as mountain are used in different locations and by different groups, approaches to automatically generating macro-maps in space and time and, finally, techniques allowing regions to be characterised and compared based on the terms used to describe them. The use of all these methods in combination allows us to come closer to a meaningful representation of place in the sense of human geography within the context of Geographic Information Science. However, our approaches focus on the naming of places and their material or perceivable properties, and there is still much work to do to properly represent place, and particularly sense of place. Nonetheless, we suggest that such approaches have considerable potential for those working in the digital humanities, and especially those concerned with contributing to a spatial turn therein.

DOI: https://doi.org/10.3366/ijhac.2015.0139

FROM SPACE TO PLACE: PLACE-BASED EXPLORATIONS OF TEXT

ROSS S. PURVES AND CURDIN DERUNGS

Abstract *New data sources, for example in the form of geotagged image libraries and digitised archives of historical text documents, provide us with new opportunities for exploring how place is described. Using a framework derived from work in human geography and information science, we illustrate how there is more to place than names and coordinates. Through a set of case studies we explore different aspects of the seemingly trivial query 'mountains in the Alps' addressing a range of issues including ambiguity, the use of vernacular names, ways in which concepts such as mountain are used in different locations and by different groups, approaches to automatically generating macro-maps in space and time and, finally, techniques allowing regions to be characterised and compared based on the terms used to describe them. The use of all these methods in combination allows us to come closer to a meaningful representation of place in the sense of human geography within the context of Geographic Information Science. However, our approaches focus on the naming of places and their material or perceivable properties, and there is still much work to do to properly represent place, and particularly sense of place. Nonetheless, we suggest that such approaches have considerable potential for those working in the digital humanities, and especially those concerned with contributing to a spatial turn therein.*

Keywords: sense of place; vernacular place names; locale; macro-maps; GIS

International Journal of Humanities and Arts Computing 9.1 (2015): 74–94
DOI: 10.3366/ijhac.2015.0139

I. INTRODUCTION

In this paper, we aim to describe our ongoing work which sets out to explore notions of place using digital sources ranging from user generated content, for example Flickr (www.flickr.com) or Geograph (www.geograph.org.uk), to digitised historical documents such as Text+Berg[1]. Underlying our research is the observation that GIScience has largely failed to deal with place, as elegantly stated by Fisher and Unwin[2]:

> Geographic Information theory articulates the idea of **absolute Euclidean spaces** quite well, but the **socially-produced and continuously changing notion of place** has to date proved elusive to digital description except, perhaps, through photography and film.

Following Fisher and Unwin, we wish to move from simply representing Euclidean spaces using traditional topographic data collected by mapping agencies for a given purpose, towards associating fluid and socially-produced conceptualisations of place with given sets of coordinates. We assert that the increasing volume of data available digitally provides us with previously unimaginable opportunities to carry out such research. In particular, in this paper we concentrate not only through the 'modern' (e.g. tagged images such as those uploaded to Flickr) but also the 'old' (in the form of digitised archives of textual documents). Central to all of our work is the idea that we can somehow relate descriptions of place to locations in space – that is to say we can locate in Euclidean space the data with which we work.

From this overarching context we resolve the following set of general research questions, which aim to be sufficiently generic such that they can make a contribution to the spatial humanities in general:

- What are the **opportunities** of our approaches and what **new insights** - on the notion of place - do we generate?
- What are the **major challenges** and the **limitations** of such methods?
- How can **theoretical context** influence the choice of methods?

In the following, we will first briefly set out some background material, defining place and its characteristics from a number of perspectives. Having done so, we will introduce some of the methodological approaches taken in our research to capturing place-based information. Finally, we illustrate the use of these methods in a number of case studies, showing how these approaches start to facilitate more nuanced exploration. In doing so, it is important to state from the outset the limitations of our work. We have primarily analysed text either as an unstructured bag of words, or through simple co-occurrence. Such approaches are typically based on frequencies, often normalised against some expected values. These methods are far from capable of representing the full breadth of

place – nevertheless, we claim they are a useful step where the spatial humanities can, and should, play a significant role.

2. BACKGROUND

Place is a familiar, if contested concept in human geography[3]. Tuan[4] elaborates on the distinction between space and place by referring to a discussion between two physicists, namely Heisenberg and Bohr. Heisenberg states:

> Isn't it strange how this castle [Kronberg] changes as soon as one imagines that Hamlet lived here? As scientists we believe that a castle consists only of stones, and admire the way the architect put them together. The stones, the green roof with its patina, the wood carvings in the church, constitute the whole castle. None of this should be changed by the fact that Hamlet lived here, and yet it is changed completely.

Ironically, the two natural scientists explore the meaning of a place by referring to characteristics that are clearly beyond physical matter or laws. Spaces, Tuan argues, are demarcated by boundaries and consist of physical matter. When we are in a strange part of town, what stretches ahead of us is space:

> In time we know a few landmarks and the routes connecting them. Eventually what was strange town and unknown space becomes familiar place.[5].

The idea that with emerging local knowledge space can be transformed into place, is also represented in Basso's[6] linguistic anthropology:

> And so, unavoidably, senses of place also partake of cultures, of shared bodies of 'local knowledge' (the phrase is Clifford Geertz's) with which persons and whole communities render their places meaningful and endow them with social importance.

Doreen Massey, in her essay 'A global sense of place'[7], zooms out from the meaning of place to individuals and at local scales, by drawing a complex picture of the world. Every place is attached meaning through its relations not only with local places, but with all other places around the globe:

> It is a sense of place, an understanding of 'its character', which can only be constructed by linking that place to places beyond. A progressive sense of place would recognize that, without being threatened by it. What we need, it seems to me, is a global sense of the local, a global sense of place.[8]

From this short overview on how space and place are represented in the humanities, it is obvious that any representation of place must consider not only the physical objects found in a space, but also the meanings which those objects have for individuals and communities. Furthermore, place is, at least according to Massey, not only a local concept, but also a global one. In contrast, in Geographic Information Science place is often used as a shorthand for a set of coordinates or, perhaps, a toponym (a *placename*). We would argue, however, that the name of the castle inhabited by Hamlet tells us little in isolation about either the material form of the space or its socially produced meaning[9].

In what follows we set out to anchor our explorations of place around Agnew's[10] definitions of places as *meaningful locations* consisting of three key aspects:

- **Location** is defined as a location in space which can be named and thus, at least implicitly, assigned coordinates.
- **Locales** are effectively affordances[11], that is to say the properties of a location which allow a particular activity to take place. Locales need not be tied to a fixed location, for example a bus or train allows particular sets of activities and interactions, in the same way as a church or indeed a mountain.
- **Sense of place** relates to the ways in which we might (or indeed might not) identify with a place and is necessarily based on our associations and experiences, or lack thereof, of a place.

It is important to note that Agnew's model is debated, but we argue that it provides us with a basic framework to move beyond the much criticised, and still predominant, reduction of place to a name and a set of coordinates (i.e. the location). Interestingly, Agnew's model of place has many overlaps with perspectives developed independently in information science to explore image classification and querying. The Panofsky-Shatford facet matrix[12] describes the *Who*, *What*, *Where* and *When* facets, each in terms of three modes, the *Specific Of*, the *Generic Of* and the *About*. A closer reading of Shatford's work, and examples of its application, reveals that the *Where* facet and its modes closely parallel Agnew's model of place. Since the information science literature often uses this matrix as a starting point in classifications, it is a useful point of reference when looking for implicitly place-centered approaches to analysis of digital content.

What then are the implications of such models? We choose to explore these by examining a potential search engine query 'Mountains in the Alps' which at first glance appears trivial, and analysing this query from a place-based

perspective. The query contains both specific and generic elements according to the Panofsky-Shatford matrix or, according to Agnew, both location(s) and locale(s). Further, its meaning is modified according to the individual who posed it, thus incorporating notions of sense of place.

Mountain is one of the most common terms offered as an example of a geographic object in empirical experiments[13]. Mountains are an excellent exemplar of locale, a place affording particular activities and interactions. This first query term is one which has seen much attention in GIScience[14] with respect to a wide range of notions, including both semantic and spatial vagueness. What is a mountain, and do we all agree? It seems obvious that this is not the case, and that a tourist from the Netherlands might have a rather different view of what is (or isn't) a mountain than a resident of Switzerland, but how can we even start to model this? Equally, how far down do we have to descend from the summit of the Matterhorn until we are no longer on its slopes? Again, it seems trivial to state that at no point one can simply 'step off the mountain', and yet typical GIS representation of objects as either points or polygons imply exactly such a property.

Moving to 'Alps', a number of issues become apparent. Firstly, are the Alps referred to in New Zealand or Europe? This is a classic example of referent ambiguity, to which most toponyms are subject[15]. Furthermore, many toponyms are also subject to semantic ambiguity[16] – that is to say does Alps refer to the European mountain range, or 'an area of green pasture on a mountainside' (Oxford dictionary)? Furthermore, having disambiguated the term Alps and identified an unambiguous referent for the term, with what region should it be associated? Not only is this region vague, but its extent might vary between different groups of individuals[17]. Finally, many toponyms are used vernacularly and are not found in official toponym gazetteers[18], as is indeed the case for Alpen (the Alps) in SwissNames, the gazetteer of Switzerland's national mapping agency.

All of these issues relate rather obviously to notions of location, and illustrate that the seemingly trivial assignment of a name to a point or region is in fact much more complex, and includes multiple challenges. Furthermore, this naming might also relate to our sense of place, in so far as our experience of a location has much to do with the region to which we might consider it to belong. Finally, the spatial relationship 'in the' seems to be a straightforward topological containment. But how can we deal with topological containment with respect to not only a vague region (the Alps), but also a concept which is in itself vague (mountains).

It is clear that to deal with the challenges posed by such a query, which we argue relate to many of the notions of place summarised in this paper, new approaches are required which take advantage of multi-faceted perspectives

Table 1. Case studies discussed in this paper.

specific / location	**1) Delineating vernacular regions** How can large digital data sources be used to delineate spatial footprints for regions that are in vernacular use and thus (often) not contained in official gazetteers (e.g. *Alps*)? **2) Drawing macro-maps from historic text** How can we automatically draw map representations of a historic text corpus containing landscape descriptions?
generic / locale	**3) Mapping locale** How can geographic locales, such as *mountains*, be linked to space? **4) Describing locale through co-occurrence** How can georeferenced text descriptions be used to gather information on locales and how can this information be structured?
specific + generic / location + locale	**5) Linking locations and locale** How can toponyms in text (i.e. locations) be associated with local knowledge (i.e. locale), such that we for instance learn about the relation between the location *Alps* and its descriptions in terms of associated locales, such as *mountain*?

relating space to place. One strand of geographic research that has considerable potential for generating such (local) information, takes advantage of so called user generated content (UGC)[19] or volunteered geographic information (VGI)[20]. A further opportunity lies in the ever increasing volume of digitised historic data, currently being embraced in the digital humanities[21], and the opportunities presented by a spatial turn with respect to such data[22].

We suggest that the trend of using multiple, identifiably human, data sources has the potential to facilitate the transformation of GIScience from a space to a place based science. Thus, we forecast that a future GIS will represent Kronberg castle in a more holistic way, where space and place based characteristics are consolidated. In this paper we set out to illustrate how we use the combination of methods from GIScience in combination with both UGC and digitised text archives, allowing us to start to explore many of the issues discussed in the context of 'Mountains in the Alps' through the five case studies (Table 1). These are each associated with notions of location and locale and, implicitly, with sense of place.

3. METHODOLOGICAL APPROACHES AND EXAMPLES

The following examples aim to illustrate approaches which address different ideas related to place, primarily in terms of location (e.g. in work with toponyms) and locale (e.g. through exploring the properties of a region). They are all taken from previous work by the authors and colleagues, but applied to new data and examples related to the query 'Mountains in the Alps' as previously discussed.

3.1 Delineating Vernacular Regions

One key task in the place-based description of locations, is adding toponyms to gazetteers which are used in a vernacular sense and not contained in official gazetteers. Furthermore, since such toponyms often refer to vague regions, methods are required which take cognisance of this property, and ideally allow reasoning with multiple vernacular toponyms or mixtures of vernacular and vague toponyms. Early work in this area focussed on identifying such toponyms through readings of text and exploring the regions associated with them through empirical experiments[23]. The development of internet search engines and associated APIs enabled the use of co-occurrence of candidate vernacular toponyms with known toponyms to identify associated regions[24]. These methods required the application of geoparsing methods to identify and disambiguate known toponyms found in text.

The advent of georeferenced UGC where, for example, images are assigned not only descriptive tags but also coordinates, allowed simpler methods to be developed, which rely upon tags in such content referring to the location under investigation, and assume that the density of usage typically correlates with the region of interest[25]. Such densities also start to allow us to represent vagueness, though the question of how to use this in a query is still non-trivial. A simple schematic of the approach taken in this paper is given in Figure 1.

However, though such methods have been illustrated by a number of authors, the availability of gazetteers containing vernacular toponyms, and particularly which allow multiple representations of the same location is still uncommon, and although there is recognition that toponyms are inherently vague and contested in gazetteer development[26], this has not filtered through to many gazetteer implementations. Equally, the identification of new, unknown vernacular toponyms is not addressed in many of these approaches, though the work of Rattenbury and Naaman[27] starts to suggest possible approaches to finding locally salient labels, which may in turn represent toponyms.

In Figure 2 we illustrate the identification of such regions for our query term (Alps) and a second mountainous region (the Cairngorms) in Scotland. In both cases we identified all Flickr images associated with a georeference and

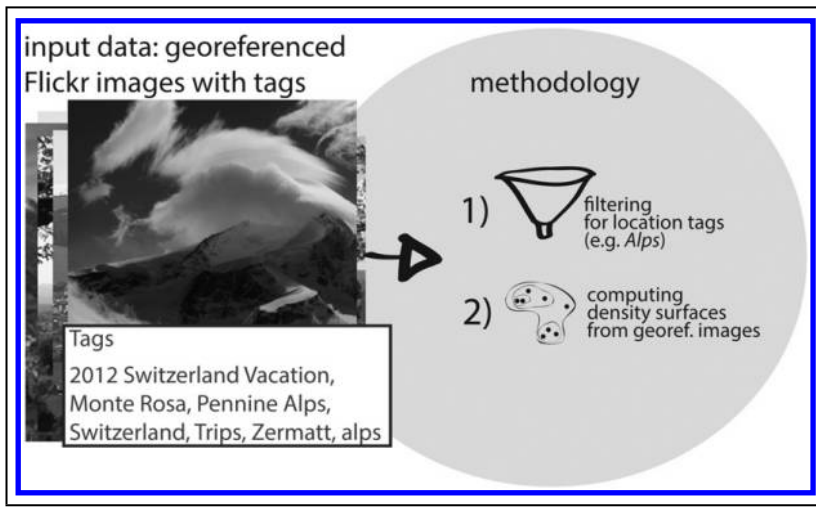


Figure 1. Methodology for delineating vernacular regions.

the corresponding tags, before using Kernel Density Estimation to estimate the regions associated with these tags. By filtering users (for example distinguishing between locals and tourists[28]) it would be possible to explore how perception of such regions varied for different groups. These regions could then form a basis for queries (for instance for texts describing visits to the Cairngorms or Alps) which might also be weighted such that descriptions in regions of highest densities were treated as being more relevant. Of course, it is important to note that the regions thus derived reflect where people take pictures, what they take pictures of, and the individuals who take pictures and upload them. Such regions are thus, as with all UGC, not representative of some global population, but rather the user groups and biases inherent in the content used to derive them[29]. Nonetheless, such data provide us with important insights into the use of vernacular toponyms describing vague regions which were previously only accessible through empirical experiments which are not scalable to large numbers of toponyms, and are also of course subject to biases.

3.2 Drawing Macro-Maps from Historic Text

Cooper and Gregory[30] introduced the term macro-maps for an approach which effectively sought to summarise the spatial content of textual archives, arguing that:

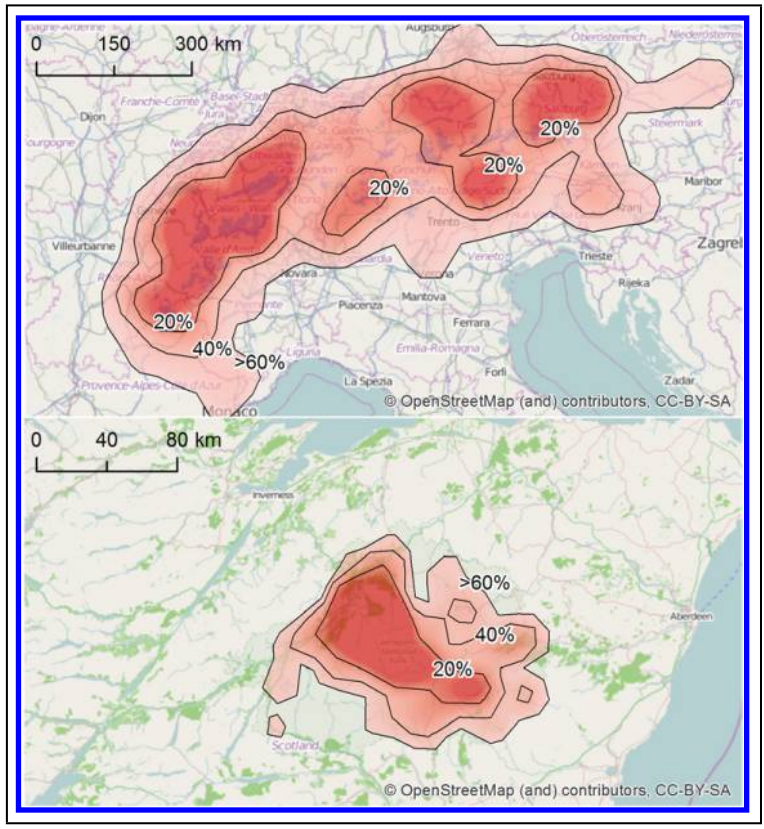


Figure 2. Kernel density surfaces representing the use of the tags Alps (top) and Cairngorms (bottom) in Flickr images of Switzerland and Great Britain. Percentages indicate quantile (e.g. top 20%) of surface volume enclosed.

> This macro-mapping will allow the literary critic to identify large-scale spatial patterns and to visualise how the representation of particular places and spaces may have shifted over time.

However, a major stumbling block in such approaches is the development of methods which automatically process large volumes of digital text, and effectively deal with the issues of toponym ambiguity previously presented. In Cooper and Gregory's work, although the archives were digitised, the process of toponym disambiguation was manual. Figure 3 describes a workflow for automatically creating a macro-map from a digitised archive.

Typical approaches to toponym disambiguation rely on rule-based methods using information found in gazetteers (e.g. population or attribute type) or

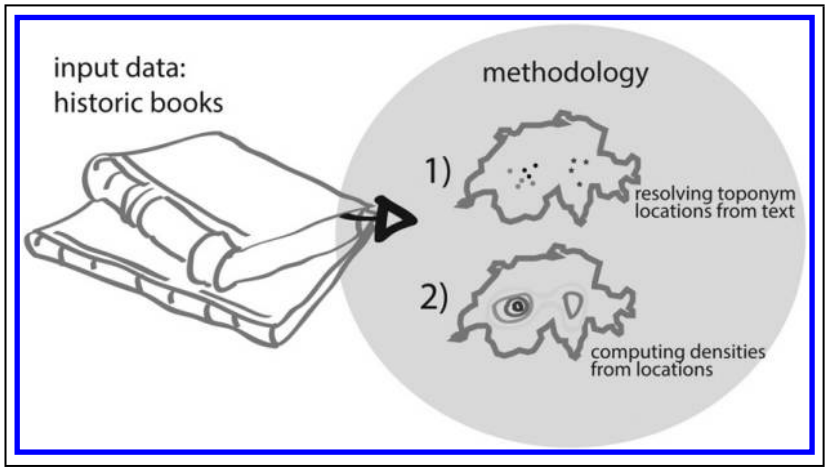


Figure 3. Methodology for drawing macro-maps from historic text .

machine-learning approaches which require labelled training data[31]. However, when disambiguating toponyms describing mountain landscapes, very little such information is available, and we developed a method which can effectively be considered to take the locale, in the form of the topographic setting of a toponym, into account[32]. This method first identified unambiguous toponym referents, before iteratively disambiguating other referents in a text using a combination of Euclidean distance and topographic similarity derived from digital elevation models (DEMs) to describe the region around a toponym in terms of slope, curvature and roughness[33].

Figure 4 shows the results of such a process, mapping the contents of the Text+Berg archive of Swiss Alpine Yearbooks firstly into space, and then into space and time (Figure 5). These macro-maps allow us to rapidly explore locations that have been the subject of discourse in Swiss mountaineering over the last 150 years. However, these maps, as presented, only allow us to explore the locations which have been discussed – effectively to automatically produce macro-maps based on toponym occurrences in documents – but not to describe the properties of these regions – for example in terms of locale. The following sections address such issues.

3.3 Mapping Locale

The advent of large volumes of UGC which are also explicitly assigned geographic coordinates gives rise to intriguing new opportunities for investigating notions of locale. In early work we explored the extent to which

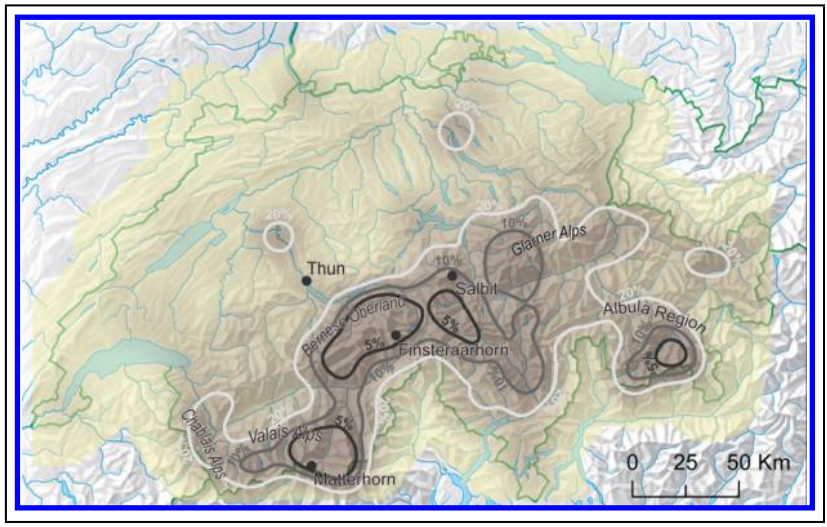


Figure 4. Macro-map of 10,000 descriptions from the Text+Berg corpus. Percentages indicate quantile (e.g. top 5%) of surface volume enclosed. Note concentrations of texts in the Bernese Oberland, Valais Alps and Albula Region.

category norms or basic levels[34] identified in traditional empirical experiments were replicated in such content[35]. Having found that such data did indeed show considerable overlaps, with similar classes of landscape feature, such as river, hill and valley, being common, we set out to explore how such landscape features occurred in space.

As was the case when mapping vernacular regions, the most straightforward approach is to use geotagged images containing these generic terms *sensu* Shatford[36] and to map these into space. However, simply mapping term densities without considering the overall image density leaves us at risk of understanding only where images are taken (and thus where people go). For this reason, we prefer to map χ-surfaces (Figure 6), highlighting regions where terms are over/underrepresented with respect to the overall distribution of images[37]. Figure 7 shows such representations for four generic terms related to locale: mountain; hill; lake and waterfall in Switzerland and Great Britain.

A number of features quickly become apparent. Firstly, mountains in Switzerland are found in broadly similar regions to those associated with the Alps in Figure 2. Secondly, lakes and waterfalls in Switzerland are punctually located, and typically (though not always) occur in areas related to mountains. Finally, the use of the term hill in Switzerland is related to only a few rather specific locations. By contrast, in Great Britain all four terms are more or less

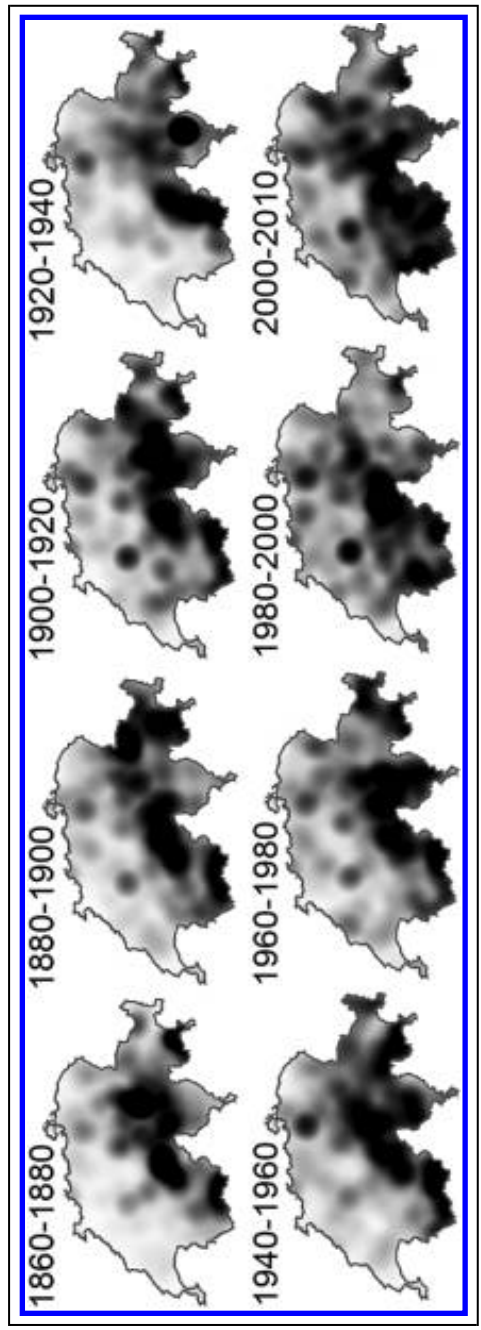


Figure 5. Temporal macro-maps of the Text+Berg corpus, computed for twenty year periods. Note spread of text to east in the period from 1880–1900, as the Rhaetian Railway opens up the region.

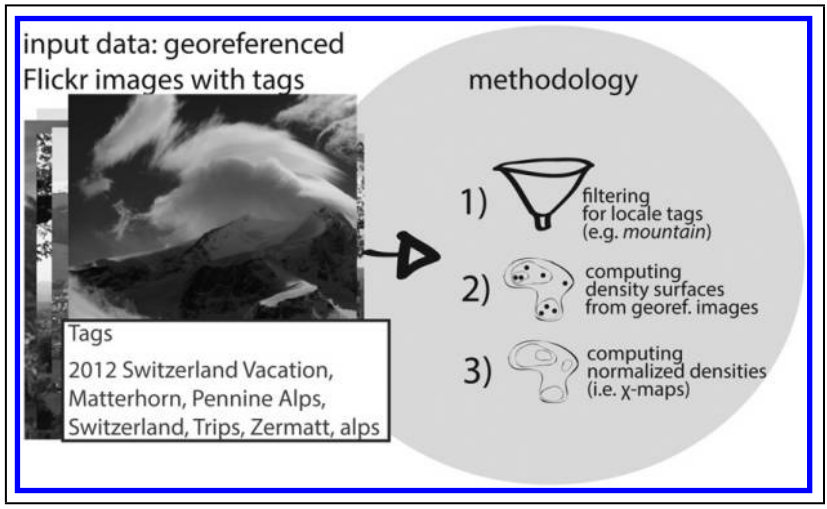


Figure 6. Methodology for mapping locale.

spatially autocorrelated. In particular, mountain and hill are found in very similar regions, and waterfalls tend to be photographed in these areas of high relief, with some exceptions (for example the extreme south-west where waterfalls appear to occur more prominently). Lake is very predominately associated with the Lake District (roughly in the centre of the landmass), indicating semantic ambiguity and the link between the region's name and its properties in this special case (the Lake District is of course so named because of the prominence of water bodies).

3.4 Describing Locale through Co-Occurrence

Having identified which terms are used where, a second challenge is to explore how locales are characterised. Through the use of treemaps[38], such co-occurrence relationships can be explored both in terms of locales, their qualities and associated activities[39]. The methodology for computing such treemaps is sketched in Figure 8.

We originally set out to automatically identify nouns, adjectives and verbs and use these as terms and their descriptors. However, our experience has shown that such automatic classifications are often not very illuminating, and we have typically relied on annotating lists of common terms as elements, qualities and activities[40] which appear to capture key parts of landscape scenes[41].

These methods allow us to start to focus on the locale in and of itself – what is it about a particular region that makes it of interest – what do people notice there and how do they use it? The treemaps (Figure 9) reveal some basic differences

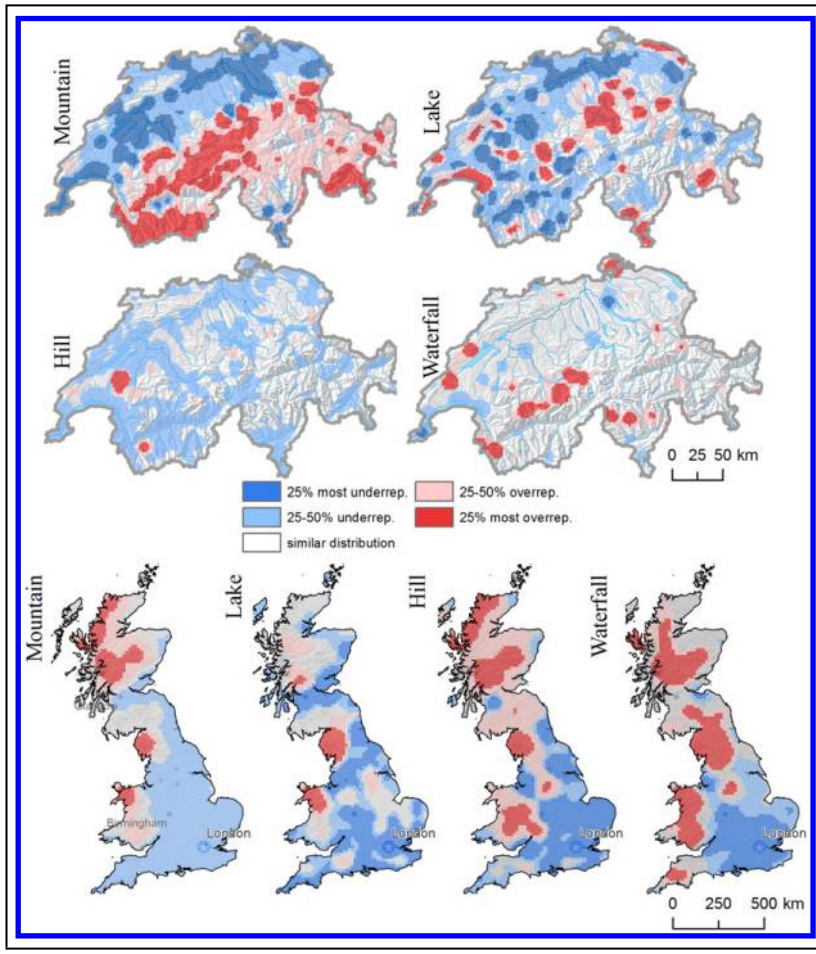


Figure 7. χ-maps for four terms associated with locale in Switzerland and Great Britain based on georeferenced Flickr images tagged with those terms. Blue regions show where terms are underrepresented with respect to the overall distribution of Flickr images, red regions indicate where terms are overrepresented.

in the use of terms. Firstly, mountain is much more prominent in Switzerland than hill, whilst in the UK hill and mountain appear to occur with similar frequencies. A closer look at the qualities co-occurring with these terms reveals further differences. Hill in the UK is associated with very similar qualities to mountain (e.g. country, rural, wild, nature) suggesting that the use of mountain and hill is more or less synonymous. However, in Switzerland hill is associated

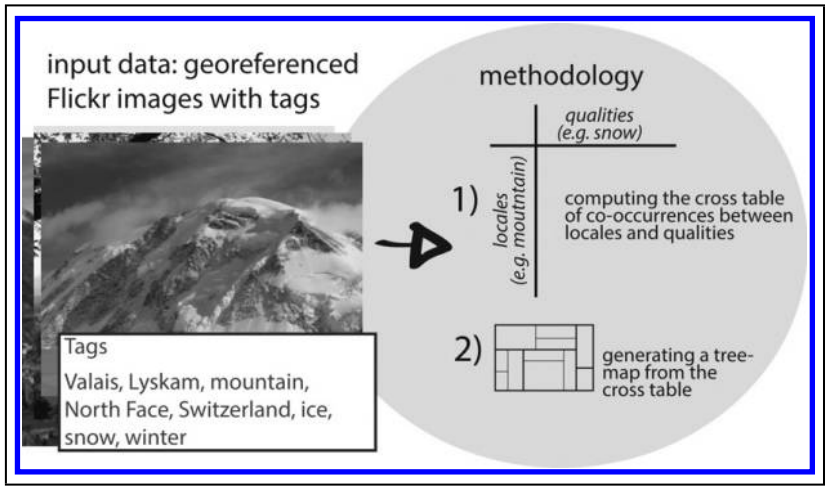


Figure 8. Methodology for describing locale through co-occurrence.

with some different qualities, in particular medieval, suggesting a quite different use (as also suggested in Figure 7).

3.5 Linking Locations and Locale

In our final example we illustrate how such explorations of locale as were illustrated above can be carried out for the rich text corpus for which we produced a macro-map in §3.2[42].

Figure 10 describes the most important methodological steps for doing so. We index content to locations in a contiguous fashion, by use of an adaptive grid, which tessellates a whole region according to the amount of content found in individual grid cells. Associated with each cell are landscape terms, for which term vectors of both raw frequencies within a grid cell (ignoring overall term occurrence) and spatial *tf-idf*[43] values can be calculated. These term vectors in turn allow us to explore the similarity of locations (in terms of cells which can be associated with toponyms) through locale (represented in this case as the prominence of particular landscape elements at a location). Figure 11 shows such term vectors for twelve locations in Switzerland.

These methods go beyond simple density based macro-mapping, and start to suggest ways in which regions can be described using locale and, potentially, sense of place. Although our current approaches treat the corpus as a bag of words, and thus mask different perspectives, it is easy to see how similar approaches could be used to describe locale from, for example, the perspective of a botanist or a mountaineer.

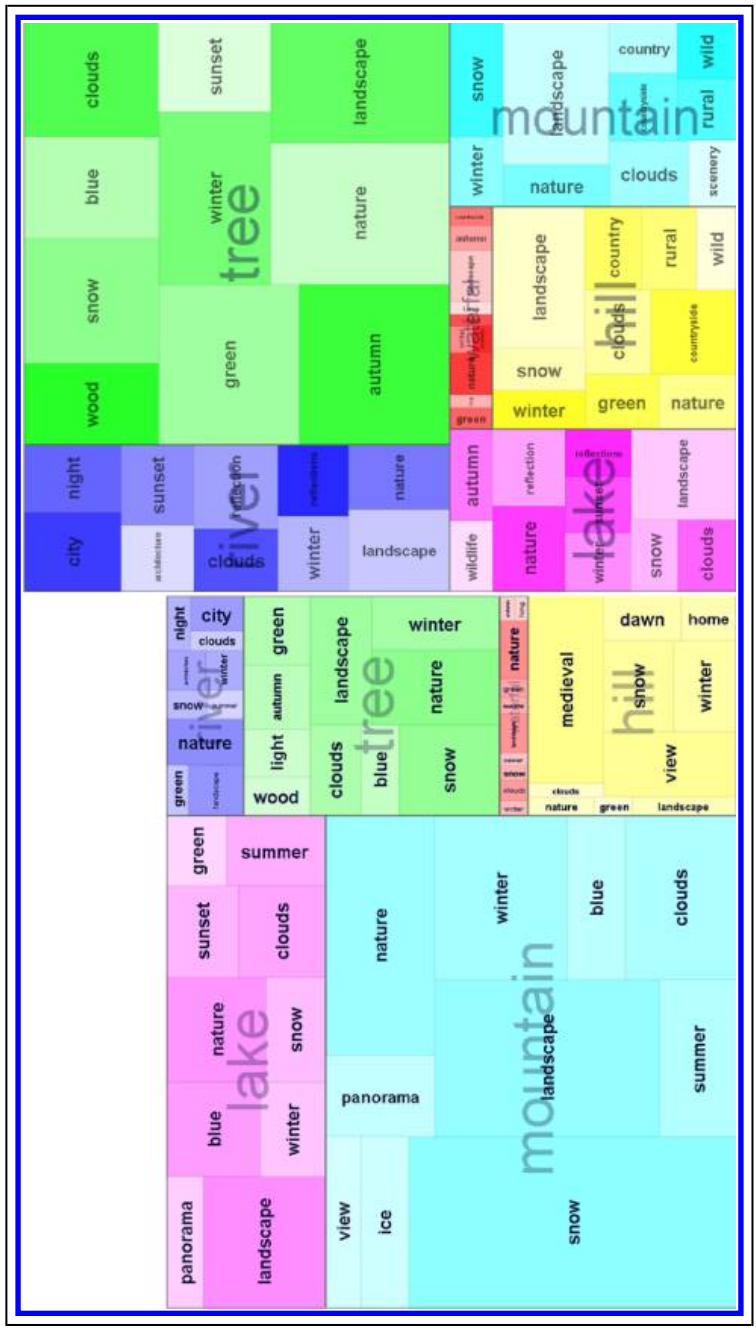


Figure 9. Treemaps showing co-occurrence of *elements* identified in Flickr (hill; mountain; lake; waterfall; river) with the ten *qualities* most commonly associated with the element for Switzerland (left) and Great Britain (right).

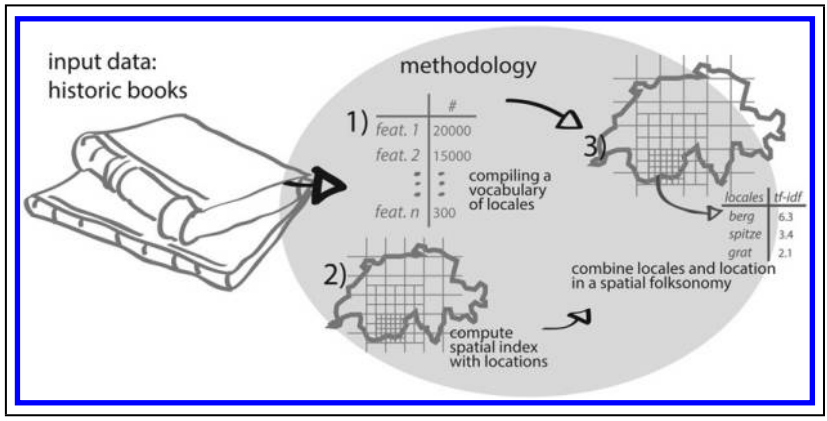
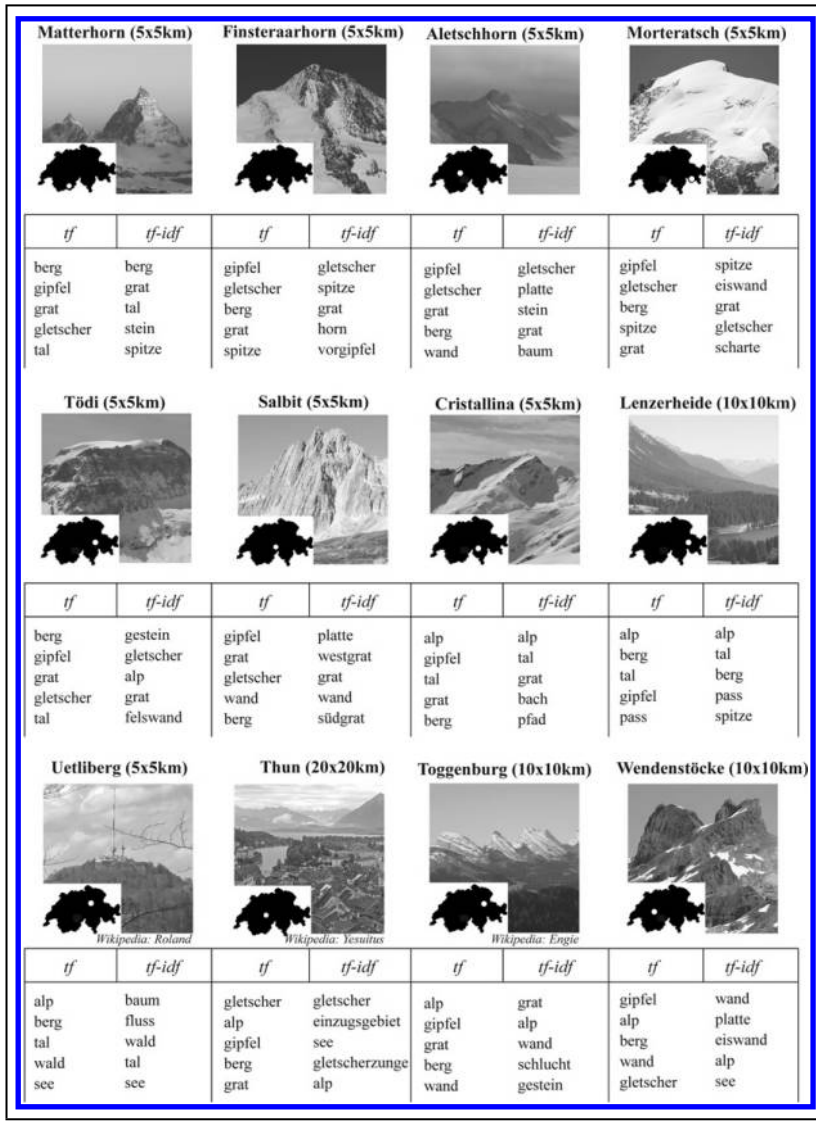


Figure 10. Methodology for linking locations and locale.

4. CONCLUDING DISCUSSION

In this paper we have tried to describe our work on both user generated content, typically in the form of geotagged images, and more traditional digitised text archives. We have concentrated on our research efforts, and tried to illustrate the overarching framework within which we are working – namely making a contribution from a GIScience perspective to representations of place, and to describe some of the methods we are using to do this. We are of course acutely aware that in so doing we have illustrated a small subset of all that is going on in this, necessarily, interdisciplinary field. Furthermore, the astute reader will have noted that though we introduced the notion of *sense of place*, we have failed to return to it in any substantive way.

So, given these limitations, what contributions have we made to the research questions we sketched in the introduction? Firstly, we believe that our paper illustrates very clearly the **opportunities** afforded to us by the combination of new and emerging data sources, which allow us to start to represent different conceptualisations of the same place, with theoretical underpinnings from, for example geography and information science. Secondly, we believe that our methods demonstrate clearly that place is more than an empty placeholder, but rather a fundamentally useful way of structuring information, which can in turn start to address some of the criticisms of GIS advanced by some of its strongest critics[44]. Thus, for example, by using locale to explore locations we gained **insights** into the way in which the concept of hill appears to mean something quite different in Great Britain to Switzerland, with its use being effectively both spatially autocorrelated and synonymous to mountain in Great Britain, while in Switzerland this is not the case.



tf	*tf-idf*	*tf*	*tf-idf*	*tf*	*tf-idf*	*tf*	*tf-idf*
berg	berg	gipfel	gletscher	gipfel	gletscher	gipfel	spitze
gipfel	grat	gletscher	spitze	gletscher	platte	gletscher	eiswand
grat	tal	berg	grat	grat	stein	berg	grat
gletscher	stein	grat	horn	berg	grat	spitze	gletscher
tal	spitze	spitze	vorgipfel	wand	baum	grat	scharte

tf	*tf-idf*	*tf*	*tf-idf*	*tf*	*tf-idf*	*tf*	*tf-idf*
berg	gestein	gipfel	platte	alp	alp	alp	alp
gipfel	gletscher	grat	westgrat	gipfel	tal	berg	tal
grat	alp	gletscher	grat	tal	grat	tal	berg
gletscher	grat	wand	wand	grat	bach	gipfel	pass
tal	felswand	berg	südgrat	berg	pfad	pass	spitze

tf	*tf-idf*	*tf*	*tf-idf*	*tf*	*tf-idf*	*tf*	*tf-idf*
alp	baum	gletscher	gletscher	alp	grat	gipfel	wand
berg	fluss	alp	einzugsgebiet	gipfel	alp	alp	platte
tal	wald	gipfel	see	grat	wand	berg	eiswand
wald	tal	berg	gletscherzunge	berg	schlucht	wand	alp
see	see	grat	alp	wand	gestein	gletscher	see

Figure 11. Locale based descriptions through landscape features for twelve locations in Switzerland. Note similarities between term vectors ranked using raw frequencies (*tf*), whilst rankings based on *tf-idf* values show more localised variation.

There are of course many **challenges** and **limitations** to the approaches we have described. Even very large digital archives can be prone to issues of bias and contributor inequality, and these problems are often hard to disentangle from meaningful signal. Thus, for example, we know little about the editorial policies of the Swiss Alpine Club over the last 150 years, and only a close reading of the texts, and associated materials could reveal to what extent our results are influenced by such effects.

In this paper we have sought to illustrate how the **theoretical context** put forward by human geographers, and in particular Agnew[45], has many parallels with practically useful approaches taken by information scientists[46]. This enabled us to build a useful framework, through which we were able to explore both location and locale in the context of mountain landscapes in what we assert is a meaningful, and potentially useful, way. However, taking account of not only local, but global ideas of place[47] will require the development of methods which not only allow us to drill down into data, but also make meaningful links between locations and locales. Finally, much of the work here introduces methods which are explored in a given setting – however, the next, and crucial, step is to use these methods in research questions defined by domain experts. We look forward to stimulating interactions with colleagues in the future.

ACKNOWLEDGEMENTS

This paper brings together methods developed with a wide range of colleagues over the years. They include Jason Dykes, Alastair Edwardes, Christian Gschwend, Chris Jones, Livia Hollenstein and Jo Wood. Thanks are due to all of them, though any errors are of course our own. Oliver Zihler is thanked for his assistance with data preparation for some of the figures herein. Comments from two anonymous reviewers helped us to clarify many points in the paper. Some of the research reported in this paper was supported by the project FolkOnt2 supported by the Swiss National Science Foundation under contract 146157 and the University Research Priority Program *Language and Space*.

END NOTES

[1] M. Volk, 'How many mountains are there in Switzerland? Explorations of the SwissTopo name list', *Search. Answers Festschrift Honour Michael Hess Occas. His 60th Birthd.,* (2009) 127–140.

[2] P. Fisher, and D. Unwin, 'Re-presenting geographical information systems', *Re-Presenting GIS*, London Wiley (London, 2005). Cited at 6.

[3] T. Cresswell, *Place – A short introduction*, (Chichester, 2011).

[4] Y. Tuan, *Topophilia. A study of environmental perception, attitudes, and values*, (New York, 1974).

[5] Tuan, *Topophilia. A study of environmental perception, attitudes, and values*, 199.

[6] K.H. Basso, *Wisdom sits in places: Landscape and language among the Western Apache*, (New Mexico, 1996). Cited at xiv.

[7] D. Massey, 'A global sense of place', *Marx Today*, 35 (1991), 24–29.

[8] Massey, 'A global sense of place', 29.

[9] for instance in R. Coates, 'Properhood', *Language*, 82 (2006), 356–382.

[10] J. Agnew, 'Space and place', *SAGE Handb. Geogr. Knowl.*, (2011), 316–330.

[11] J. J Gibson, *The ecological approach to visual perception*, (London, 1979).

[12] S. Shatford, 'Analyzing the subject of a picture: A theoretical approach', *Cat. Classif. Q.*, 6 (1986), 39–62.

[13] B. Smith, and D. M. Mark, 'Geographical categories: an ontological investigation', *Int. J. Geogr. Inf. Sci.*, 15 (2001), 591–612.

[14] two examples are: B. Smith, and D. M. Mark, 'Do mountains exist? Towards an ontology of landforms', *Environ. Plan. B Plan. Des.*, 30 (2003), 411–427; P. Fisher, J. Wood, T. Cheng, 'Where is Helvellyn? Fuzziness of multi-scale landscape morphometry', *Trans. Inst. Br. Geogr.*, 29 (2004), 106–128.

[15] E. Amitay, N. Har'El, R. Sivan, A. Soffer, 'Web-a-Where: Geotagging web content', in M. Sanderson, K. Järvelin, J. Allan, P. Bruza, eds., *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval.* (New York, 2004), 273–280; D. Buscaldi, 'Approaches to disambiguation of toponyms', in R. S. Purves, and C. B. Jones, eds., *Lett. Geogr. Inf. Retr.*, (2011), 16–20.

[16] Amitay, 'Web-a-Where: Geotagging web content'.

[17] C. B. Jones, R. S. Purves, P. D. Clough, H. Joho, 'Modelling vague places with knowledge from the Web', *Int. J. Geogr. Inf. Sci.*, 22 (2008), 1045–1065.

[18] C. Davies, I. Holt, J. Green, L. Diamond, 'User needs and implications for modelling vague named places', *Spat. Cogn. Comput.*, 9 (2009), 174–194.

[19] for instance M. Haklay, and P. Weber, 'Openstreetmap: User-generated street maps', *Pervasive Comput. IEEE*, 7 (2008), 12–18.

[20] M. F. Goodchild, 'Citizens as sensors: the world of volunteered geography', *GeoJournal*, 69 (2007), 211–221.

[21] D. M. Berry, *Understanding digital humanities*, (New York, 2012).

[22] G. Lock, 'Representations of space and place in the humanities', in D. J. Bodenhamer, J. Corrigan, T. Harris, eds., *Spat. Humanit. Futur. Humanit. Sch.*, (Bloomington, 2010), 89–108.

[23] for instance D. Montello, M. Goodchild, J. Gottsegen, P. Fohl, 'Where's downtown?: Behavioral methods for determining referents of vague spatial queries', *Spat. Cogn. Comput.*, 3 (2003), 185–204.

[24] for instance Jones et al. 'Modelling vague places with knowledge from the Web'.

[25] C. Keßler, P. Maué, J. T. Heuer, T. Bartoschek, 'Bottom-up gazetteers: Learning from the implicit semantics of geotags', in K. Janowicz, M. Raubal, S. Levashkin, eds., *GeoSpatial Semant.*, (2009), 83–102; C. Grothe and J. Schaab, 'Automated footprint generation from geotags with kernel density estimation and support vector machines', *Spat. Cogn. Comput.*, 9 (2009), 195–211; L. Hollenstein and R. S. Purves, 'Exploring place through user-generated content: Using Flickr to describe city cores', *J. Spat. Inf. Sci.*, 1 (2010), 21–48.

[26] L. L. Hill, Georeferencing: *The geographic associations of information*, (Cambridge, 2009); H. Southall, R. Mostern, M. Berman, 'On historical gazetteers', *Int. J. Humanit. Arts Comput.*, 5 (2011), 127–145.

[27] T. Rattenbury and M. Naaman, 'Methods for extracting place semantics from Flickr tags', *ACM Trans. WebI*, 3 (2009), 1–30.

[28] F. Girardin, F. Calabrese, F. D. Fiore, C. Ratti, J. Blat, 'Digital footprinting: Uncovering tourists with user-generated content', *Pervasive Comput. IEEEI*, 7 (2008), 36–43.

[29] L. Hollenstein and R. S. Purves, 'Exploring place through user-generated content: Using Flickr to describe city cores', *J. Spat. Inf. Sci.*, 1 (2010), 21–48.

[30] D. Cooper and I. N. Gregory, 'Mapping the English Lake District: a literary GIS', *Trans. Inst. Br. Geogr*., 36 (2011), 89–108. Cited at 105.

[31] B. Martins, I. Anastácio, P. A. Calado, 'Machine learning approach for resolving place references in text'*, Mach. Learn*., (2010), 221–236.

[32] C. Derungs and R. S. Purves, 'From text to landscape: Locating, identifying and mapping the use of landscape features in a Swiss alpine corpus', *Int. J. Geogr. Inf. Sci*., (2013).

[33] for instance J. Iwahashi and R. Pike, 'Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature', *Geomorphology*, 86 (2007), 409–440.

[34] B. Tversky and K. Hemenway, 'Categories of environmental scenes', *Cogn. Psychol*. 15 (1983), 121–149.

[35] A. Edwardes and R. S. Purves, 'A theoretical grounding for semantic descriptions of place', *Web Wirel. Geogr. Inf. Syst*., (2007), 106–120.

[36] Shatford, 'Analyzing the subject of a picture: A theoretical approach'.

[37] J. Dykes, R. S. Purves, A. Edwardes, J. Wood, 'Exploring Volunteered Geographic Information to describe place: Visualization of the "Geograph British Isles" collection', *Proceedings of GIS Research*, (2007), 256–267; C. Gschwend and R. S. Purves, 'Exploring geomorphometry through User Generated Content: Comparing an unsupervised geomorphometric classification with terms attached to georeferenced images in Great Britain', *Trans. GIS*, 16 (2012), 499–522.

[38] J. Wood and J. Dykes, 'Spatially ordered treemaps', *IEEE Trans. Vis. Comput. Graph*., 14 (2008), 1348–55.

[39] Tversky and Hemenway, 'Categories of environmental scenes'.

[40] three examples are: Derungs and Purves, 'From text to landscape: Locating, identifying and mapping the use of landscape features in a Swiss alpine corpus'; A. Edwardes and R. S. Purves, 'A theoretical grounding for semantic descriptions of place', *Web Wirel. Geogr. Inf. Syst.*, (2007), 106–120; R. S. Purves, A. J. Edwardes, J. Wood, 'Describing place through user generated content', *First Monday*, 16 (2011).

[41] Tversky and Hemenway, 'Categories of environmental scenes'.

[42] Derungs and Purves, 'From text to landscape: Locating, identifying and mapping the use of landscape features in a Swiss alpine corpus'.

[43] tf-idf (term frequency – inverse document frequency) is a standard measure in information retrieval which takes into account both the frequency of a term and its overall rate of occurrence. It is broadly analogous to the use of χ-values as shown in Figure 7.

[44] J. Pickles, *Ground truth: The social implications of geographic information systems*, (New York, 1994).

[45] Agnew, 'Space and place'.

[46] Shatford, 'Analyzing the subject of a picture: A theoretical approach'.

[47] Massey, 'A global sense of place'.